

DETERMINING NATURAL SCALES OF ECOLOGICAL SYSTEMS

REBECCA L. HABEEB,¹ JESSICA TREBILCO,¹ SIMON WOTHERSPOON,² AND CRAIG R. JOHNSON^{1,3}

¹*School of Zoology and Tasmanian Aquaculture and Fisheries Institute, University of Tasmania, Private Bag 5, Hobart, Tasmania 7001 Australia*

²*School of Mathematics and Physics, University of Tasmania, Private Bag 37, Hobart, Tasmania 7001 Australia*

Abstract. A key issue in ecology is to identify the appropriate scale(s) at which to observe trends in ecosystem behavior. The characteristic length scale (CLS) is a natural scale of a system at which the underlying deterministic dynamics are most clearly observed. Any approach to estimating CLSs of a natural system must be able to accommodate complex nonlinear dynamics and must have realistic requirements for data. Here, we compare the robustness of two methods to estimate CLSs of dynamical systems, both of which use attractor reconstruction to account for the complex oscillatory dynamics of ecological systems. We apply these techniques to estimate CLSs of spatial multispecies systems of varying complexity, and show that both methods are robust for the simplest system, but as model complexity increases, the Pascual and Levin metric is more robust than that of Keeling et al. Both methods demonstrate some sensitivity to the choice of species used in the analysis, with closely connected species producing more similar CLSs than loosely connected species. In this context, connectivity is determined both by the topology of the interaction network and spatial organization in the system. Notably, systems showing complex spatial self-organization can yield multiple CLSs, with larger length scales indicating the emergent dynamics of interactions between patches. While the prediction r^2 metric of Pascual and Levin is suitable to estimate CLSs of complex systems, their method is not suitable to apply to most real ecosystems because of the requirement of long time series for attractor reconstruction. We offer two alternatives, both based on prediction r^2 , but where repetition in space is largely (the “short time series” method) or wholly (the “sliding window” method) substituted for repetition in time in attractor reconstruction. Both methods, and in particular the short time series based on only three or four sequential observations of a system, are robust in detecting the primary length scale of complex systems. We conclude that the modified techniques are suitable for application to natural systems. Thus they offer, for the first time, an opportunity to estimate natural scales of real ecosystems, providing objectivity in important decisions about scaling in ecology.

Key words: attractor reconstruction; characteristic length scale; community dynamics; ecosystem; nonlinear dynamics; spatial and temporal dynamics; spatial scale; spatiotemporal models.

INTRODUCTION

A fundamental goal in ecology is to determine the dynamical processes underlying observed patterns. The single greatest difficulty confronting this important objective is that any pattern detected, and ultimately the understanding of the underlying dynamical processes, depends on the spatial scale at which we make our observations (Wiens 1989, Levin 1992, 2000, Schneider 1994). The challenge is to identify the appropriate scales of observation for ecological investigation (Levin 1992). Are there natural or “characteristic” scales of ecological systems that are optimal for observing a system’s behavior that provide unambiguous information about the dynamic, and which minimize noise in the signal measured? If so, then the capacity to identify these scales may be useful to applied ecology, par-

ticularly to monitor and detect meaningful change in ecosystem state (Rand 1994, Bishop et al. 2002).

The search for a means to identify natural or “characteristic” length scales (hereafter CLSs) in ecological systems stems back at least to the 1950s (Grieg-Smith 1952, Kershaw 1957), but several more recent attempts have also addressed the problem (e.g., Carlile et al. 1989, De Roos et al. 1991, Schneider 1994). Most approaches have assumed either that ecological systems are stationary in space and time, that fluctuations are random around a stationary global average (e.g., Rand and Wilson 1995), or that any trends detected are linear (see Turner et al. 1991). These kinds of approaches fail to take into account the dynamical nature of ecological systems in space and time, which is arguably characterized by nonlinear oscillatory behaviors (e.g., Hastings et al. 1993, Ellner and Turchin 1995, Sole and Bascompte 1995, Little et al. 1996, Pascual and Ellner 2000).

In contrast, the relatively recent techniques of Keeling et al. (1997) and Pascual and Levin (1999) to iden-

Manuscript received 9 September 2004; revised 22 February 2005; accepted 25 February 2005; final version received 5 April 2005. Corresponding Editor: A. M. Ellison.

³ Corresponding author. E-mail: craig.johnson@utas.edu.au

tify CLSs are exceptional in that they both accommodate the complex nonlinear oscillations and non-uniform patterns in spatial variance that are pervasive in real ecosystems. The crucial development of these new approaches is the application of attractor reconstruction (Takens 1981) and prediction algorithms (Kaplan and Glass 1995) from nonlinear time series analysis to characterize dynamics at particular scales of observation. Takens (1981) shows that the attractor describing the dynamics of an entire system can be estimated from the time series of any one species. Both the Keeling et al. (1997) and Pascual and Levin (1999) techniques seek to identify the spatial scale that best distinguishes the deterministic dynamic, or trend, in a system from noise. The underlying tenet is that if the scale used to sample an ecosystem is too small, observations are swamped by noise due to strong correlations among individual samples (Durrett and Levin 2000, Wilson and Keeling 2000). If the sampling scale is too large, the nontrivial dynamics will be averaged out because distant parts of the landscape begin to act independently (Keeling et al. 1997, Pascual and Levin 1999, Wilson and Keeling 2000). The CLS is an intermediate scale which most clearly reflects the underlying deterministic signal (Pascual and Levin 1999, Wilson and Keeling 2000; for illustration, see Appendix A).

While the techniques of both Keeling et al. (1997) and Pascual and Levin (1999) show promise, thus far they have only been applied to simple model systems. Furthermore, these techniques require unrealistically long time series so that their application to most real ecosystems is impractical. Thus, if the goal of developing a technique to identify CLSs in real ecosystems is to be realized, there are two critical steps. First, these techniques must be able to indicate unambiguous length scales for dynamical systems more complex than those investigated to date. Second, the technique of attractor reconstruction needs to be modified to obviate the need for long time series.

These challenges define the two broad aims of our research: to assess the robustness of the approaches of Keeling et al. (1997) and Pascual and Levin (1999) so that we can be confident in applying the techniques to complex ecological systems, and to develop a method of attractor reconstruction that does not require an unrealistically long time series. We assess robustness by examining the capacity of the two techniques to identify unambiguous length scales for model systems spanning a spectrum of complexity, and by determining the sensitivity of CLS estimates to (1) the initial spatial arrangement of individuals in a system, (2) the choice of species used in attractor reconstruction, and (3) parameter choices in attractor reconstruction. We then develop two alternative methods which rely wholly or partially on substituting space for time in attractor reconstruction. One approach is based on using short time series obtained from multiple locations in space, while

in the other we slide "windows of observation" through space at a single point in time, entirely substituting repetition in space for repetition in time.

METHODS

Our overall approach was first to determine whether the metrics of Keeling et al. (1997) and Pascual and Levin (1999) indicate unambiguous length scales for systems of varying complexity across a range of species and choices of parameters for attractor reconstruction. We then derived two alternative approaches to estimate CLSs based on the most robust metric (i.e., that of Pascual and Levin 1999), where we substitute space for time to remove the need for long time series. Finally, we compared the estimates of CLSs derived using our two new methods with that of the original approach based on long time series.

Existing methods for estimating CLSs of dynamic oscillating systems

Both the methods of Keeling et al. (1997) and Pascual and Levin (1999) are based on a sequence of time series, each sampled at a different spatial scale. The time series for scale l is constructed by marking an $l \times l$ window on the landscape and recording the abundance or density of a single species over time as the system evolves. Nonlinear time series methods are used to make predictions for the time series, and the accuracy of these predictions is estimated. This process is repeated for a range of scales and prediction accuracy is compared across all scales.

Both Keeling et al. (1997) and Pascual and Levin (1999) use attractor reconstruction from nonlinear time series analysis as the means of prediction (Takens 1981, Kaplan and Glass 1995, Little et al. 1996). Assuming there is a deterministic system underlying the dynamics, observations of a single variable in the system over time are used as a proxy from which to predict the dynamics of the whole system. For example, in a predator-prey system, prey (or resource) density can be measured and used to predict the deterministic signal of the entire system (Rand and Wilson 1995). The strengths in using attractor reconstruction for prediction of ecological systems are that there are no assumptions of steady state, and that complex fluctuations can be accommodated. A more detailed description of the method of attractor reconstruction is given in Appendix B, while more comprehensive treatments can be found in several introductory texts on nonlinear dynamics and attractor reconstruction (Kaplan and Glass 1995, Abarbanel 1996). In summary, the phase space of the ecosystem dynamic is built from time delay coordinates of a single species, which act as surrogates for the unobserved variables of the system (Casdagli 1989, Abarbanel 1996, Kantz and Schreiber 1997). The trajectories in phase space are then used to reconstruct a set that is topologically equivalent to the attractor of

the full system (Farmer 1982, Sugihara et al. 1990, Kaplan and Glass 1995).

While both approaches for CLS estimation use attractor reconstruction as their basis, they differ in their measures of prediction error. Keeling et al. (1997) apply attractor prediction to calculate deviations around the underlying deterministic behavior and then to plot error variance (termed “error X ”) as a function of window size, producing a variance spectrum. For sufficiently large windows, the relative variance initially increases at a rate proportional to window size, and then plateaus. The window size at which error X reaches the asymptote is the length scale where correlations have decayed and values at different sites behave like independent random variables. This scale, where the windows become statistically independent, is identified as the CLS. It defines the “window size” in which the full spatial dynamics of the system can be observed.

Pascual and Levin’s (1999) method is a variant of the approach of Keeling et al. (1997) and aims to extract the scale where the ratio of determinism to noise is maximized. This CLS will usually be slightly smaller than that required for the onset of independence defined by Keeling et al. (1997), because it is the minimum window size where the dynamics of the system can be accurately predicted (Pascual and Levin 1999). At each window size, the degree of determinism is evaluated from the prediction accuracy of the algorithm derived from attractor reconstruction. Pascual and Levin (1999) then examine how the prediction error of the trajectories changes with spatial scale (Kaplan and Glass 1995). They plot a statistic termed the prediction r^2 (or degree of determinism) for each window size to produce a variance spectrum similar to that of error X . The scale where the prediction r^2 spectrum first attains an asymptote with respect to window size identifies the scale of maximum determinism, or the CLS of the system.

By definition, sampling at a scale smaller than the CLS is suboptimal, and as window size increases towards the CLS, prediction accuracy should increase substantially. Sampling at scales larger than the CLS offers comparatively little gain in accuracy. Thus, the CLS is estimated as the scale or window size at which prediction accuracy plateaus.

Spatial models

Models of varying complexity, in terms of species richness, the network topology defining the number of direct and indirect interspecific interactions in the system (Johnson and Seinen 2002), and spatial pattern, were developed using the COMPETE software (see Supplement), which is a probabilistic individual-based system to model spatial competition between sessile colonial organisms. The models follow the fate of competing individuals in a two-dimensional landscape, and can demonstrate complex behaviors indicating nonlinear dynamics and spatial self-organization (Johnson

1997, Johnson and Seinen 2002). Any network topology among S species is possible, including intransitive loops (e.g., where $S_1 > S_2$, $S_2 > S_3$, $S_3 > S_1$; with $S_x > S_y$ indicating that species x outcompetes and displaces species y), which arise commonly in benthic marine systems (Johnson and Seinen 2002). We used models with intransitivities in their network topology because they enable persistence stability (sensu Johnson and Mann 1988) of the system without the need for elaborate model closures and forcing functions. We update the system synchronously, and use periodic (toroidal) boundary conditions.

Four model systems were implemented (in order of complexity of spatial pattern): symmetric networks of three, six, and 12 species, and a model of the dynamics of the benthos of a coral reef (Fig. 1). The three-species system is the simplest intransitive loop as described above; i.e., a circular network with binary interaction outcomes, so that each interaction has one unambiguous winner and one loser (Fig. 1A). The six-species system involves a symmetrical network in which each species overgrows and is overgrown by two species ($S_1 > (S_2, S_3)$, $S_2 > (S_3, S_4)$, \dots , $S_6 > (S_1, S_2)$), with all other interactions as standoffs. In this system, the species spatially organize into two distinct groups of three species and, if the model is run for sufficient time, either group may eventually dominate (Fig. 1B). The 12-species system has a network structure of ($S_1 > (S_2, S_3, S_4)$, $S_2 > (S_3, S_4, S_5)$, \dots , $S_{12} > (S_1, S_2, S_3)$) and organizes into three groups of four species, any of which may begin to dominate, as the emergent dynamic unfolds (Fig. 1C). In all three models the growth rates of all species are identical. For simplicity, there was no disturbance or mortality, and no recruitment of propagules to unoccupied sites. In the six- and 12-species models, two scales of self-organization are emergent: that of the colony and that of the patch, which is a distinctive group of colonies of several species. While the identity of species in particular patches is consistent, there is oscillation in the areal dominance of patches over long time series (1000x time steps). Patch dominance varies depending on the initial random configuration of “recruits” at time step zero and on stochasticity in execution of the rules of local interactions. This dominance arises as a result of finite landscape size, and becomes less obvious with very large landscapes (i.e., larger than $\sim 500 \times 500$ cells). Colonies tend to become more aggregated over long time series, resulting in larger multispecies patches.

The coral reef model is more complex than the other three because (i) parts of colonies or whole colonies may die with subsequent recruitment to unoccupied space, and (ii) competitive outcomes are not binary so that for all pairwise interactions among species S_x and S_y , the probability of $S_x > S_y$ and $S_y > S_x$ is nonzero. There are 12 physiognomic life forms in the model (Fig. 1D). Neighbor-specific growth rates, interaction outcomes, and mortality and recruitment rates have

been parameterized from direct observations of communities on the Great Barrier Reef. The emergent community composition is very similar to that of real reefs after an appropriate period without a major disturbance event (notably, similar results are also evident from models of temperate marine benthic systems; see Dunstan and Johnson 2005). In this context, the coral reef model most closely simulates a natural system. We selected three physiognomic groups for analysis in our exploration of alternative methods to identify CLSs: (i) turf and coralline algae, (ii) corymbose and digitate Acroporidae (acroporid corals), and (iii) Faviidae (faviid corals). Acroporid and faviid corals occupy less cover than the turf and coralline algae, with faviids occurring in the smallest patches.

Determining the robustness of existing techniques to estimate CLS

We considered several properties in assessing the robustness of prediction r^2 (Pascual and Levin 1999) and error X (Keeling et al. 1997) spectra derived from long time series. First, by definition, neither measure should show any dependence on window size when spatial and temporal patterns are random, i.e., in the absence of any deterministic signal or oscillatory behavior. We show that this result holds true (Appendix C).

The other aspects of robustness considered are (1) the capacity to indicate an unambiguous length scale for spatial models of a range of complexity, (2) the capacity to identify similar length scales for different runs of the same system, (3) the capacity to identify similar length scales from different species in the same system, and (4) for any one model run, sensitivity of the CLS estimate to different arbitrary choices of parameters required for attractor reconstruction, namely τ (time delay), d_e (embedding dimension), and k (number of nearest neighbors; see Appendix B for details).

The overall approach was to run all four model systems for 10 000 time steps on landscapes of 500×500 cells, sampling landscapes from times 201 to 10 000 (the first 200 time steps are ignored while the system self-structures). Each model was begun with a random spatial arrangement of "recruits" (10% total cover) on the initial landscape, with identical amounts of each species. We sampled different window sizes l ($l = 5$ to 495 in steps of 5) within the 500×500 landscape, and observed those windows through time. At each time step there was a single window of each size. The density of each species for each window size $l \times l$ was determined, generating a separate time series for each l . For a selected species, the attractor of the system in d dimensional space was estimated for each time series (see Appendix B). For each model run, we derived estimates of CLSs using methods of both Keeling et al. (1997) and Pascual and Levin (1999).

In calculating the CLSs, abundances of the selected species were scaled by window area to convert each

time series of absolute counts to a time series of density. Attractor reconstruction was undertaken as described in Appendix B for particular choices of τ , d_e , and k ($k = 10$ unless otherwise specified). The CLS was estimated as the window size, l , at which the error X or prediction r^2 spectrum reached an asymptote. Monte Carlos of 100 independent runs of each model were used to calculate mean error X and prediction r^2 spectra with 95% confidence intervals. It should be noted that while we are interested in the sensitivity and robustness of both measures, prediction r^2 will yield a slightly smaller length scale than error X for the reasons outlined earlier, so we are not concerned with whether the two techniques provide similar absolute estimates of length scales.

Robustness to model complexity.—Because prediction r^2 and error X have thus far been calculated only for simple spatial model systems, it is necessary to ascertain whether the techniques produce meaningful results for more complex systems. Variance spectra of prediction r^2 and error X were produced for the three-, six-, and 12-species models, and for the coral reef model. Spectra from the two methods were compared for each of the four models.

Robustness to initial conditions.—For each model system, we assessed the variability of the CLS estimated from the dynamics of a given species over 100 runs of the model, each with a different initial random configuration.

Robustness to choice of species.—A system's attractor is built with observations on a single species. In theory, the choice of species to reconstruct the entire system's attractor is arbitrary, as every species reflects the same underlying attractor (Abarbanel 1996). Previously, CLSs have been generated from only a single species within the system, implicitly assuming that all species will indicate the same CLS (Keeling et al. 1997, Pascual and Levin 1999). We tested this assumption for each model by comparing the CLSs estimated from different species in the same system.

Robustness to parameters of attractor reconstruction.—The estimated CLS will depend in part on the accuracy of the attractor reconstruction, which itself depends on appropriate choices of the reconstruction parameters (τ , d_e , and k , see Appendix B; Buzug and Pfister 1992, Kantz and Schreiber 1997). No single unambiguous value exists for any of these parameters for a particular reconstruction, and indeed different techniques to estimate these parameters commonly yield dissimilar values (Buzug and Pfister 1992, Schreiber 1995, 1999, Kantz and Schreiber 1997). Thus, we examined the sensitivity of the estimated CLS to a range of reasonable potential choices of these parameters, as indicated below.

1. *Time delay (τ).*—From a mathematical perspective, the choice of delay is arbitrary because the data set is assumed to be infinitely long (Kantz and Schreiber 1997). However, for a finite set of data, the choice

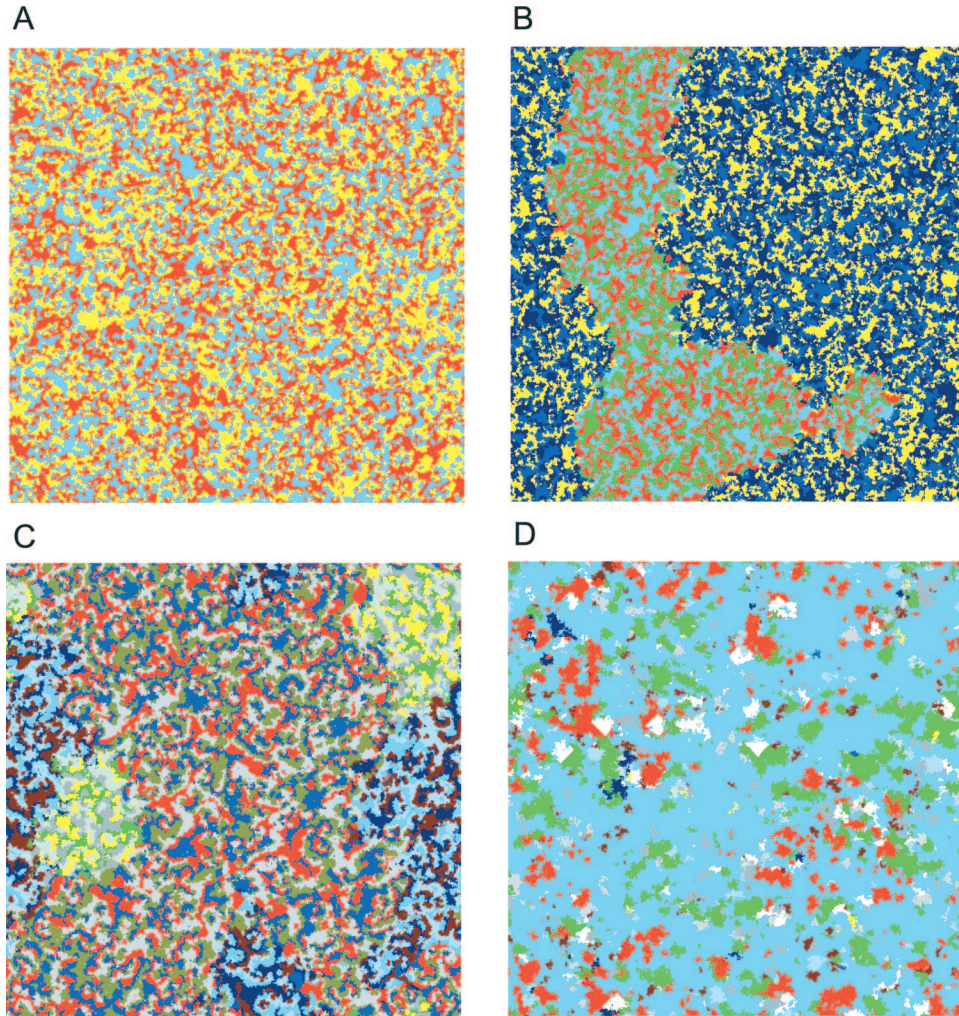


FIG. 1. Examples of 500×500 landscapes at time step 10 000 of the four model systems used to generate data: (A) three-species system, (B) six-species system, (C) 12-species system, and (D) coral reef system with 12 physiognomic groups. Each color represents a different species or group. Groups selected for analysis from the coral reef system are the dominant turf and coralline algae (light blue), digitate and corymbose *Acroporidae* (red patches), and *Faviidae* (dark gray patches). The different models represent a range of complexity of structure, dynamics, and spatial self-organizing. Note that some models self-organize at several spatial scales. All model runs used random initial configurations of recruits covering 10% of the landscape at time step zero.

of τ dictates the quality of the reconstructed trajectory (Liebert and Schuster 1989). If τ is too small, the coordinates $x(t_i)$ and $x(t_i + \tau)$ will be almost identical, offering redundant information about the state space. Alternatively, if τ is too large, the coordinates will be almost uncorrelated and their connection to one another is effectively random (Abarbanel 1996). The goal is to determine the delay where coordinates are independent while preserving their dynamical relationship (Nichols and Nichols 2001). We chose τ as the first minimum point of the mutual information (MI) function (Appendix B). However, because identifying the first minimum point from the plot of MI vs. τ requires subjective interpretation, we assessed the robustness of CLS estimates for a variety of choices of τ that might be considered reasonable using the MI approach.

2. *Embedding dimension (d_e)*.—If an attractor is projected in too few dimensions, the observed orbits will overlap and distinct segments on the attractor become confused (Abarbanel 1996). The appropriate d_e allows the attractor to be sufficiently unfolded in space such that this overlap no longer occurs. Over-embedding (embedding in too many dimensions) requires larger numbers of coordinates, and increases the likelihood of noise in the dimensions of the embedded space where no dynamics are operating (Kennel et al. 1994). The d_e is chosen as the smallest dimension required to sufficiently unfold the attractor, and is indicated as the first minimum of the false nearest neighbors vs. dimension curve (see Appendix B). We assessed the robustness of CLS estimates over a range of embedding dimensions around the value suggested by the false nearest neighbor method.

3. *Number of nearest neighbors (k).*—The number of nearest neighbors (k) used in the prediction is a tunable parameter that influences the quality of the prediction (see Appendix B). If too few neighbors are picked, then important nonrandom information may be missed. If too many are picked, the points may be widely spread in space, decreasing the accuracy of the prediction (Kantz and Schreiber 1997). We assessed the sensitivity of CLS estimates over a range of reasonable choices of k .

Alternative methods of CLS detection not dependent on long time series

To address the problem of unrealistic data requirements of the existing approaches based on analysis of long time series, we derive two alternative methods which we term the “short time series” and “sliding window” approaches. Both new methods use prediction r^2 (because this metric proved most robust), and substitute repetition in space for repetition in time. In the “short time series” method, time series are obtained over only three to four time steps, but at many locations in space. In the “sliding window” method, space is wholly substituted for time by sliding a window of observation across space at a single point in time. The logic underpinning these developments is that distant locations in space within the same dynamical system are likely to be at different points on the system’s attractor. If so, then sampling sequentially in space should produce a reconstruction sufficient to replace long time series sampling. Substituting space for time has been successful in other contexts (see, for example, Allain and Cloitre 1991, Marcos-Nikolaus et al. 2002).

Short time series analysis.—In this approach, short time series from multiple locations in space are embedded to reconstruct a system’s attractor piece-wise. Our data are from landscapes consisting of a large but finite array of contiguous cells where each cell can be occupied by a single species. We position windows so that they overlap spatially, with successive windows displaced by the width of a single cell, either horizontally or vertically (Fig. 2A). A short time series is obtained for each window. Thus, the short time series approach samples the whole landscape at every time step over a short time. The embedding parameters τ and d_c for attractor reconstruction are not determined in the same way as for long time series analysis. In principle, any value of τ should suffice for an embedding (Takens 1981), but in practice some delays are more effective (Kantz and Schreiber 1997). Here, we use a time delay of one and we ensure that the spacing between landscapes allows the system to evolve from one time step to the next. The number of embedding dimensions d_c is the number of time steps sampled minus one, as the final time step provides a data set against which the accuracy of predictions is assessed.

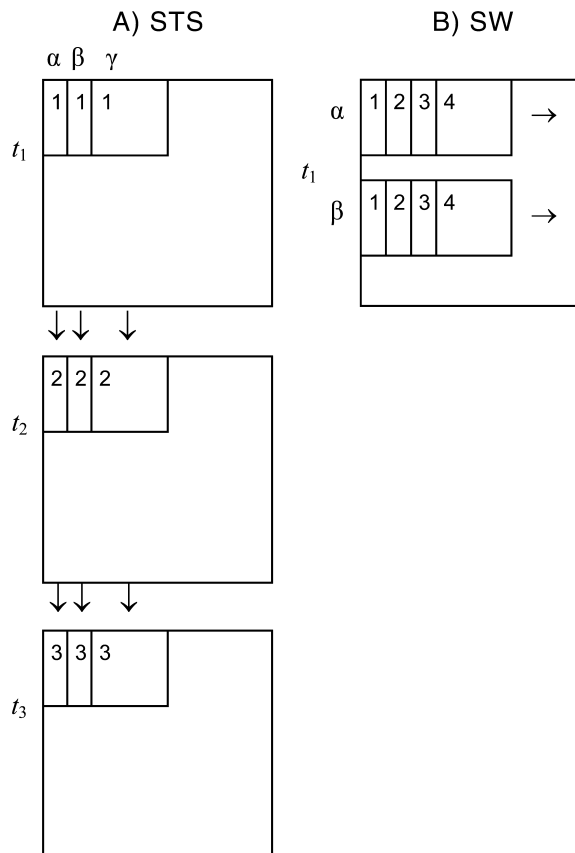


FIG. 2. Diagrams of (A) the short time series (STS) and (B) sliding window (SW) methods of attractor reconstruction. The large square represents an entire landscape that is pixelated at some resolution such that each pixel is occupied by a single species. The small squares represent windows of observation on the landscape. (A) The short time series approach uses a temporal sequence of only three or four consecutive time steps (e.g., $t_1 - t_3$). At each time step we sample over the entire landscape, with successive windows separated by a single pixel, horizontally or vertically. In this particular representation, there are three time steps t_1 , t_2 , and t_3 , and three data series at positions α , β , and γ . Three short time series are generated, and the system’s attractor is reconstructed piecewise using the delayed coordinates. (B) In the sliding window approach, multiple short data series are generated for a single time step t_1 by sliding windows horizontally across the landscape. The vertical position of each series differs by one pixel, so the entire landscape is sampled. The sequence in each data series as the window of observation slides across the landscape is indicated as 1, 2, 3, 4, . . . In this example, two data series are shown at positions α and β .

Embeddings for different values of d_c are evaluated in the results.

A modified prediction algorithm is required for attractors reconstructed from short time series. When using long time series, changing the window size does not change the number of points used to predict the attractor, because replication occurs through time. However, in short time series embeddings where replication is partially spatial, larger windows will gen-

erate fewer points in the reconstructed attractor than smaller windows, which could bias prediction accuracy. Thus, predictions for the short time series are based on a random sample of 1000 windows, which compensates for the discrepancy between small and large windows. Random samples are also used for k nearest neighbor predictions in the sliding window approach described below.

Sliding window analysis.—The sliding window sampling method also uses multiple short data series to reconstruct an attractor in sections, but there is no temporal component. A single landscape is sampled by sliding windows of observation cell by cell (or pixel by pixel) horizontally across the landscape, in a manner similar to the “gliding-box” described by Allain and Cloitre (1991) and Plotnick et al. (1996). Each data series begins one cell lower than the previous series so that the whole landscape is sampled (Fig. 2B). Attractor reconstruction from a spatial series assumes that spatial data capture the general shape of the system’s dynamics in a similar way to a time series. The assumption is reasonable if distant parts of the landscape are out of phase, that is, if they are on different parts of the system’s attractor. Because of the large model landscapes used in the current study, this assumption is met for all analyses.

There are two sensible options in selecting the spatial delay, τ , in the sliding window attractor reconstruction. One approach is to use the same delay for every window size, as for the time series embeddings. However, for spatial sampling the units of embedding delay are cells (or pixels) rather than time steps. Thus, a potential problem with using delays of a fixed number of cells is that successive embedding dimensions will have a greater overlap for large window sizes than for small window sizes, which could bias CLS estimates. An alternative is to use delays set as some proportion of the linear dimensions of the window, $\tau = \alpha \times \text{window length}$, provided that this proportion is some whole number of cells. We evaluated fixed and proportional delays by assessing their performance for the “null case,” that is, landscapes composed of independent, discrete valued pixels (results of comparison are presented in Appendix D). The results indicated that proportional delays were more appropriate, and so we adopted proportional delays in our sliding window analyses.

RESULTS

Robustness of existing techniques based on long time series

For each model system, we compare spectra based on error X and prediction r^2 . Length scales determined using error X (after Keeling et al. 1997) are designated as CLS_k , while those based on prediction r^2 spectra (after Pascual and Levin 1999) are designated as CLS_p .

Robustness of CLS estimates to initial conditions

For the simplest model system (three species), spectra of both error X and prediction r^2 were of the expected shape, increasing to an asymptote as a function of window length, and demonstrating a single length scale (Fig. 3A). As expected, CLS_k (50–60 cells) was slightly larger than CLS_p (30–40 cells; Pascual and Levin 1999). There was little variation between runs with different initial conditions. However, with increasing complexity of the model system, CLS_k became more difficult to determine and confidence intervals around the error X spectrum broadened considerably (Fig. 3B, C, D). The error X spectrum for the 12-species system was not the expected positive asymptotic shape, and was not readily interpretable (Fig. 3C). Similarly, based on error X , no length scale could be determined for the coral reef model because the curve deviated dramatically from the expected shape with no inflection or asymptote (Fig. 3D). This was true of individual spectra as well as the mean spectrum depicted in Fig. 3D. Conversely, CLS_p was not highly sensitive to changes in initial conditions for any of the models, and produced an interpretable curve for the full range of models we examined (Fig. 3).

Robustness of CLS estimates to choice of species

Estimates of CLS_k and CLS_p in the three-species model system were not dependent on species identity. All three species are ecologically equivalent in the system (Johnson and Seinen 2000), and displayed curves identical to those in Fig. 3A. In contrast, curves for the more complex model systems were often highly sensitive to the choice of species (for example, Figs. 4 and 5). CLS_k depended heavily on the species used in the attractor reconstruction for all three complex model systems. For example, in the 12-species system, CLS_k ranged from 50–60 cells for one species to 240–250 cells for another (Fig. 5A). For spectra based on error X , species in the same self-organized patch demonstrated similar curves (Figs. 4A, 5A, C). However, not only were error X spectra dissimilar for species from different patches in the same run (Figs. 4A, 5A, C), but the overall shape of the curve, and therefore the CLS_k indicated, changed markedly among runs for the same species. For example, one species in the 12-species system had a CLS_k of approximately 80–90 cells in one run (Fig. 5A), but 290–300 cells in another (Fig. 5C). The error X spectra from the coral reef model displayed little variation from run to run, but only spectra for some species were interpretable (not shown).

As model complexity increased, prediction r^2 spectra also usually separated into groups that corresponded to species within the same spatially self-organized patch. However, CLS_p estimates of the different groups of species were more similar than indicated by spectra based on error X (e.g., Fig. 4A vs. 4B). In the two most complex systems, CLS_p was more sensitive to species,

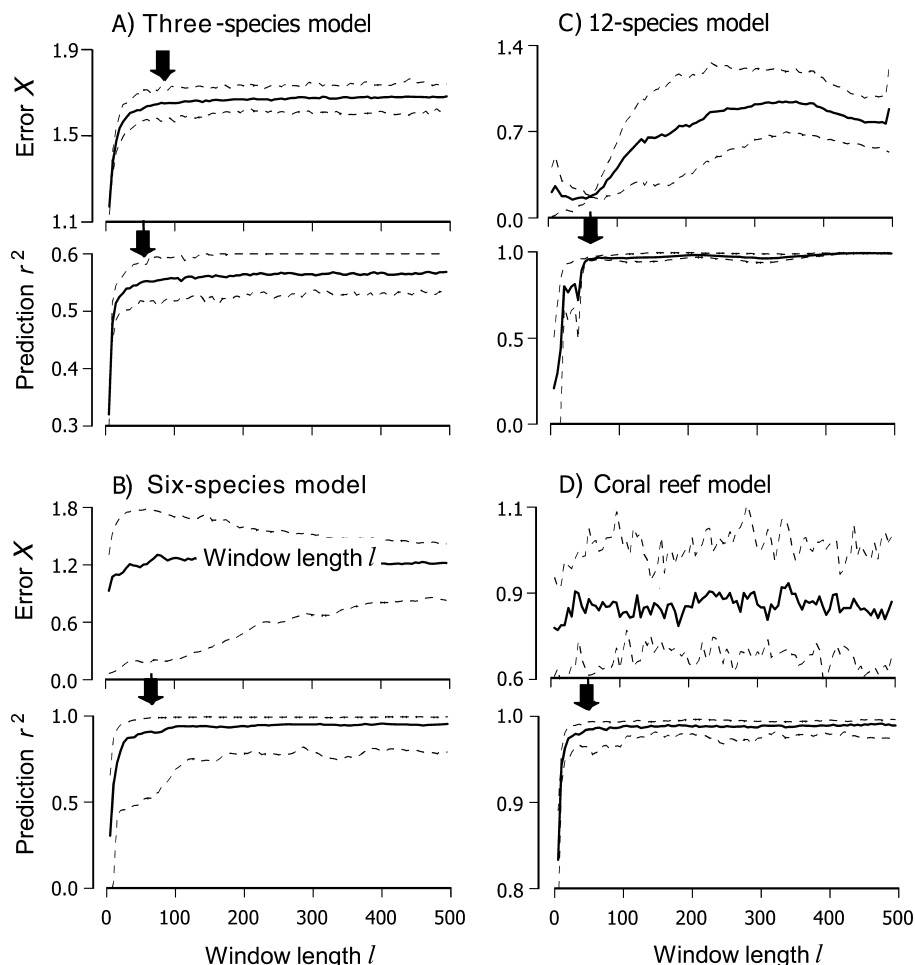


FIG. 3. Average spectra (solid lines) based on error X (Keeling et al. 1997) and prediction r^2 (Pascual and Levin 1999) and the 95% confidence intervals (dotted lines) for a selected species, calculated from a Monte Carlo of 100 independent runs for the (A) three-species model, (B) six-species model, (C) 12-species model, and (D) coral reef model. For the three-, six-, and 12-species models, the CLS was determined for a randomly selected species. The CLS for the coral reef model is based on the spatially dominant group, comprising turf and coralline algae. Estimated CLSs are indicated by arrows where curves are interpretable.

with the scale of the first asymptote differing by up to 50 cells between species groups (Fig. 5B, D). For the six- and 12-species systems, the curves of prediction r^2 for a given species sometimes demonstrated multiple peaks (Figs. 4B, 5B), indicating the potential for more than one length scale in a single system.

Robustness to choices of time delay in attractor reconstruction.—Changes in time delay of the three-species system shifted the CLS curves on the y-axis, but did not change the magnitude of the CLS estimate for either method (Fig. 6A, B). Similarly, time delay had little effect on estimates of the CLS interpreted from curves of the six-species system. However, in the two most complex model systems, CLS_k was notably less robust, with the shape of the curve changing with delay. For the coral reef model in particular, the error X curves were ambiguous at best, and varied with the delays (Fig. 6C). Conversely, CLS_p was robust to changes in

delay in all four model systems we examined (e.g., Fig. 6D).

Robustness to choices of embedding dimension in attractor reconstruction.—Changes in the embedding dimension of the three-species and six-species systems shifted the curves on the y-axis, but had no effect on estimates of CLS_k or CLS_p . In the 12-species system, the shape of the error X curve remained robust to increasing dimension, but interpreting the curves to define the CLS was difficult (Fig. 7A). For the coral reef system, the error X spectra indicated an interpretable CLS_k only with an overly large embedding dimension (Fig. 7C). For both of these more complex systems, the estimates of CLS_p were reasonably robust to changing dimension (Fig. 7B, D).

Robustness to choices of k nearest neighbors in attractor reconstruction.—The number of k nearest neighbors used to reconstruct the attractor of the three-

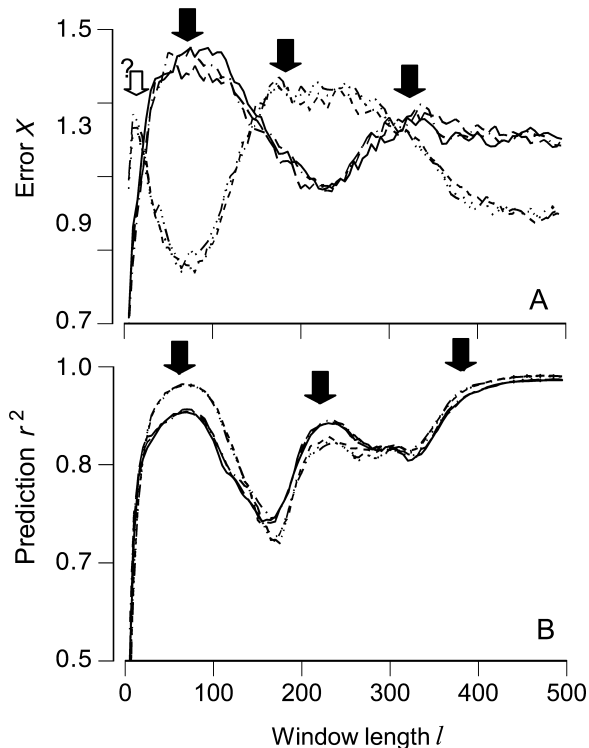


FIG. 4. Variance spectra for each species of the six-species model system based on (A) error X and (B) prediction r^2 . In this system, species spatially self-organize into two groups of three (Fig. 1B), which is reflected as two groups of curves on the graphs. Data shown are for a single run of the model. Solid arrows indicate estimated CLSs where curves are interpretable, while open arrows with question marks indicate ambiguous CLSs where interpretation is unclear. Note the potential of multiple CLSs, i.e., length scales larger than the primary (smallest) CLS.

species system had no effect on estimates of CLS_k or CLS_p . However, with the more complex systems, the shape of the error X spectra shifted from an inverted shape with low numbers of neighbors to a curve of the expected shape with excessively high numbers of neighbors (Fig. 8A). In contrast, spectra based on prediction r^2 for the same system were robust to varying numbers of neighbors in attractor reconstruction, and indicated a consistent length scale (Fig. 8B).

Alternative methods of CLS detection

Our results indicate clearly that the prediction r^2 measure (Pascual and Levin 1999) is robust in providing an unambiguous estimate of CLS_p across a range of model complexities, initial conditions, choice of species, and choice of parameters in attractor reconstruction. Because of this robustness we chose to use prediction r^2 in our alternative methods, which differ from the original approach of Pascual and Levin (1999) only in their approach to attractor reconstruction. We refer to the original method of CLS estimation based on prediction r^2 (Pascual and Levin 1999) as a "long time

series" method, which distinguishes it from the new methods we derive based on very short time series (the "short time series" method) and on a single spatially resolved landscape (the "sliding window" method).

Short time series analysis.—Our important overall result is that prediction r^2 spectra derived from attractor reconstruction based on the "short time series" method indicated unambiguous length scales for all four models we examined (Figs. 9B, C, 10C, 11B, 12D). Moreover, for any one model, the smallest CLS (i.e., primary CLS) indicated from spectra produced using both long and short time series were similar (Table 1). It is also encouraging that CLS estimates from prediction r^2 curves derived using the short time series approach are robust to the number of embedding dimensions used for attractor reconstruction (Fig. 9B, C). While results are presented for embeddings of only two and three dimensions for the three-species system, analyses using larger numbers of embedding dimensions (eight, nine, or 10 dimensions) also indicated similar CLSs (not shown). CLS estimates for the six- and 12-species systems and for the model coral reef system were also robust to the number of embedding dimensions, and spectra were consistent when the length of the time step was varied (results not presented).

Prediction r^2 curves were consistent between model runs for the three-, six- and 12-species systems indicating CLSs of 20–40 cells for the three-species system, 30–50 cells for the six-species system and 40–60 cells for the 12-species system (Figs. 9B, C, 10C, and 11B respectively). The CLS for turf and coralline algae in the coral reef model was clearly defined between 20 and 40 cells (Fig. 12D) but spectra derived from corymbose and digitate acroporid and faviid corals were more variable and not readily interpretable (Fig. 12E, F). Changing the length of the time step or changing the number of time steps used in the analysis did not affect variability in results among different runs for these coral groups which occur at relatively low cover.

Two notable differences between CLSs estimated using the long time series and short time series methods arose. First, prediction r^2 curves developed from the long time series analyses were not consistent between species, while prediction r^2 curves from short time series analyses were (for example, cf. Fig. 10A, C). Second, curves of species using long time series often indicate several peaks, suggesting the possibility of multiple length scales (for example, Fig. 4B), while only a single CLS is usually evident in spectra from the short time series. Both approaches yield a similar primary CLS, which is the smallest length scale and is indicative of local dynamics among species. We interpret additional larger length scales to indicate emergent dynamics.

A characteristic of the six-species system on small landscapes ($\sim 200 \times 200$ cells) is that over many time steps and without disturbance, the system's attractor is nonstationary. A shifting attractor is evidenced by a

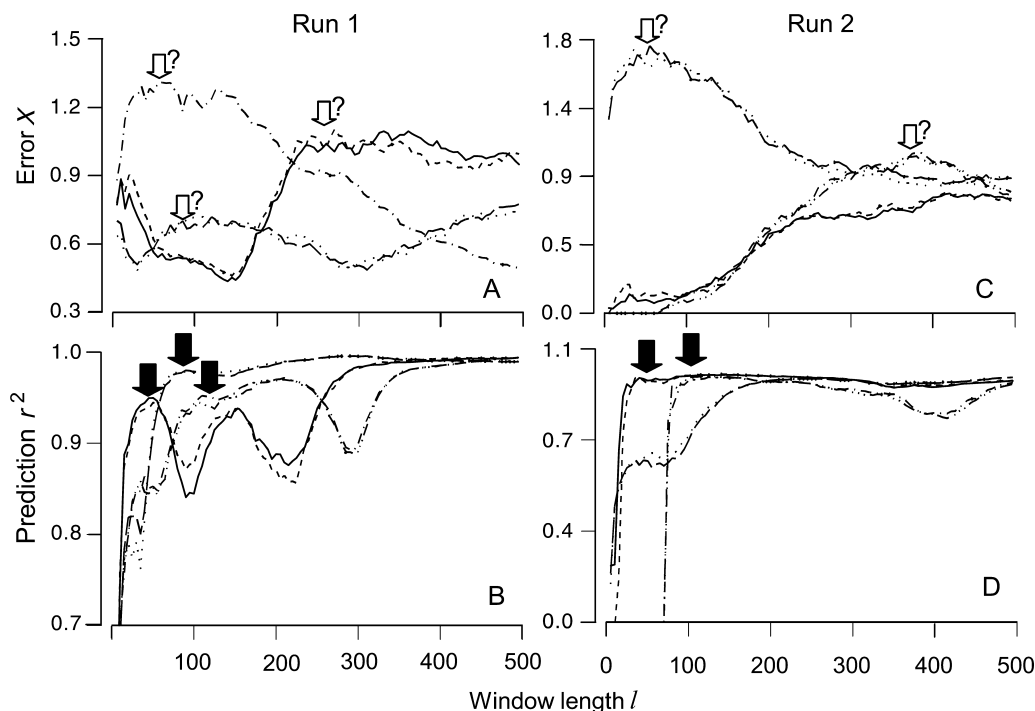


FIG. 5. Variance spectra for six species of the 12-species model system shown from two separate runs. For run 1, spectra are based on (A) error X and (B) prediction r^2 , and for run 2, spectra are given for (C) error X and (D) prediction r^2 . Species organize into three groups of four on the landscape, and species in the same patch on the landscape show similar CLS curves. Solid arrows indicate estimated CLSs where curves are interpretable, while open arrows with question marks indicate ambiguous CLSs where interpretation is unclear. For clarity, we have identified only the primary length scales for each group of species; however, multiple peaks in the curves (B) may be evidence of multiple length scales.

particular patch type expanding to eventually dominate the landscape (e.g., Fig. 13A, B), and/or coalescence of many small patches of a particular type into a single large patch of that type. The identity of the dominant patch type and the rapidity with which it realizes dominance, or the rate at which small patches of a particular type coalesce to form large patches, may vary from one run of the model to the next. In the example in Fig. 13, the community composition of the 6-species system changed gradually over 5000 time steps, so that one patch type became more abundant at the expense of the other (cf. Fig. 13A, B). Interestingly, using short time series analyses, this shift was reflected as a change in the shape of spectra, and therefore as a change in the CLS estimates derived at the beginning and at the end of the run (compare Fig. 13C, D). This kind of nonstationary behavior in the six-species system can be eliminated and the system made stationary by introducing a low level of disturbance, and allowing open recruitment of all species to disturbed areas with equal probability. As was the case for the short time series analysis (not shown), the long time series analysis of a stationary six-species time series produced almost identical prediction r^2 curves and CLS estimates for all six species (Fig. 13F). This is quite different from the non-stationary case where the spectra and CLS estimates for the two groups of three species based on

analysis of long time series were distinctly different (Fig. 13E). Unlike the six- and 12-species systems, the coral reef model is essentially stationary (after a self-structuring period of ~ 200 time steps) and so variability in CLS estimates between runs was minimal.

Sliding window analysis.—The most obvious difference between the prediction r^2 spectra produced by this method and those derived using the short and long time series approaches is that the spectrum is inverted relative to the shapes based on short and long time series analysis. Thus, the spectra yield interpretable estimates of length scales as the first minimum in the curve. Likely reasons for the inversion are addressed in *Discussion*.

In determining parameters for this method of attractor reconstruction, we have shown that delays proportional to the window size are more appropriate than are fixed delays (Appendix D). However, the question of how to select a suitable proportional delay arises. For all four model systems, the first minimum on curves of scaled mutual information vs. τ was at $\tau = 0.8 \times$ window length, suggesting that this is the most appropriate embedding delay for attractor reconstruction. However, sliding window analyses for the three-species system indicated that CLS estimates were sensitive to the choice of proportional delay (cf. Fig. 9D, E). Estimated CLS ranges increased from 20–30 cells to 50–

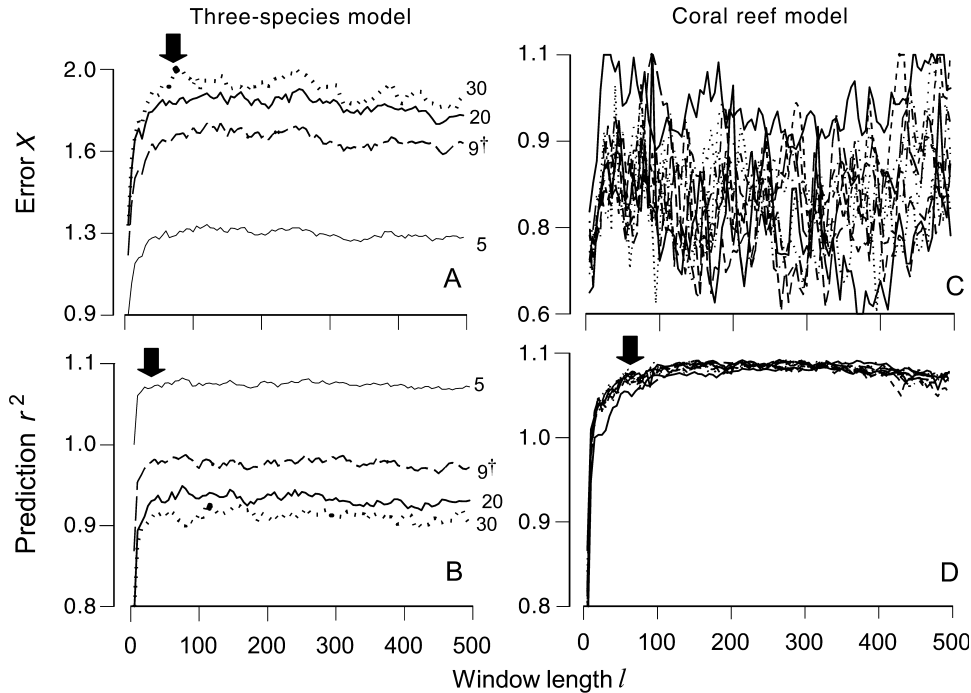


FIG. 6. The effect of choice of time delay (τ) on spectra constructed from one species of the three-species system (A and B) and the dominant species group (filamentous turf and coralline algae) of the coral reef system (C and D). In (A) and (B), the time delays are shown to the right of the relevant spectrum based on (A) error X and (B) prediction r^2 . A dagger (\dagger) denotes the delay indicated from the mutual information method. In (C) and (D), different lines denote spectra calculated from delays between 20 and 200 (in steps of 20) based on (C) error X and (D) prediction r^2 . Arrows indicate the estimated CLSs of these spectra.

60 cells when the delay changed from $\tau = 0.8 \times \text{window length}$ to $\tau = 0.2 \times \text{window length}$, respectively. Despite this sensitivity, the most easily interpretable curve was provided by using the delay indicated by the scaled mutual information (here $\tau = 0.8 \times \text{window length}$; Fig. 9E). Moreover, using this delay indicated a CLS virtually identical to that based on analysis of a long time series from the same system (cf. Fig. 9A, E).

Indeed, CLSs estimated with $\tau = 0.8 \times \text{window length}$ using the sliding window method corresponded to those estimated from analyses of long time series for all four model systems. CLS estimates for the six- and 12-species systems were between 30–40 cells and 20–30 cells respectively (Figs. 10D, 11C). All physiognomic groups in the coral reef model indicated a similar CLS range of 20–40 cells (Fig. 12G–I). However, just as was the case for the 3-species system, CLS estimates for the more complex systems were also sensitive to the proportional delay chosen. Notably, for a given value of τ , the prediction r^2 spectra were generally consistent between model runs and across species within a system.

Comparison of CLS estimates derived by all methods

For each of the four model systems that we analyzed, the three different techniques yielded similar estimates

of the smallest CLS (Table 1). CLS ranges for the six- and 12-species systems were alike across all the different methods, and were consistently larger than those for the 3-species system. Our short time series and sliding window approaches provided more precise CLS estimates for the six- and 12-species system than did the long time series method. However, prediction r^2 spectra produced using long time series analysis more often suggested multiple CLSs for a given species within the six- and 12-species systems than did the alternative analyses (Table 1).

DISCUSSION

The issue of spatial scale is a central theme in ecology. The scale at which a system is observed affects relationships between pattern and process, and between space, time, and organizational complexity (Levin 1992, 2000, Levin et al. 1997, Tilman and Kareiva 1997, Tyre et al. 1997, Dieckmann et al. 2000, Wilson and Keeling 2000, Molofsky et al. 2002). While investigators have acknowledged the need to address ecological questions at appropriate scales (Carlile et al. 1989, Wiens 1989, De Roos et al. 1991), how these “appropriate scales” are identified is often ambiguous. The application of methods from nonlinear time series analysis (Rand and Wilson 1995, Keeling et al. 1997, Pascual and Levin 1999) has refined a crucial aspect

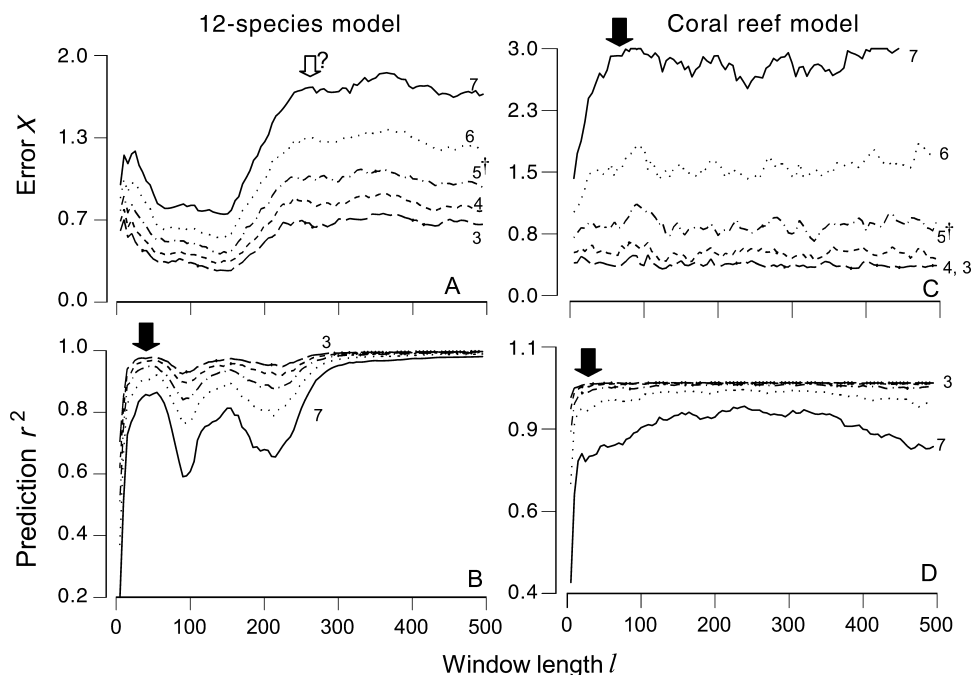


FIG. 7. The effect of choice of embedding dimension (d_e) for a single species in the 12-species system (A and B), and for the dominant species group (filamentous turf and coralline algae) of the coral reef system (C and D). Different lines denote particular embedding dimensions (shown to the right of curves) used in attractor reconstruction based on (A and C) error X and (B and D) prediction r^2 . A dagger (\dagger) denotes the dimension indicated using the false nearest neighbors method. Solid arrows indicate estimated CLSs where curves are interpretable, while open arrows with question marks indicate ambiguous CLSs where interpretation is unclear.

of the study of scale in ecology, allowing a shift in focus from observing mean behaviors to extracting the deterministic signal from dynamical systems. Towards the ultimate goal of estimating CLSs of natural systems, here we evaluated the robustness of these relatively recent methods, and then developed modifications to reduce the data requirements necessary to estimate CLSs. We first discuss the robustness of the original long time series methods, and then evaluate the general behaviors of our alternative methods in detecting the smallest or primary length scale of spatial systems. We examine the interpretation of multiple length scales, as indicated by different species groups or by a single species in the system, and finally, we comment briefly on the possible future of CLSs in ecology.

Robustness of estimates

Both the methods of Keeling et al. (1997) and Pascual and Levin (1999) maintained high robustness and low sensitivity to parameter choices in the simplest model system. As expected (Pascual and Levin 1999), CLS_k was consistently larger than CLS_p , reflecting the need for larger “windows of observation” to observe the full spatial dynamic than to accurately predict trends in dynamics. However, as the complexity of the model systems increased, the error X spectra of Keeling et al. (1997) became more difficult to interpret, and

spectra for the 12-species and coral reef model systems were not interpretable (Fig. 3C, D). Also, with increasing model complexity, error X spectra, and therefore estimates of CLS_k , were increasingly sensitive to changes in initial conditions, to the species on which the analysis focused, and to values of parameters used in attractor reconstruction. The reason for the ambiguity and, in some cases, failure (Fig. 6C) of error X spectra to indicate a CLS is unclear. Error X may be more susceptible to random noise than prediction r^2 or, alternatively, may be more sensitive to a system’s complexity and particular dynamics than is prediction r^2 . However, even if the observed sensitivity of these spectra are accurate reflections of subtle features of a particular dynamic, this level of sensitivity is not helpful towards the overall objective, which is to identify the CLS of a system irrespective of the particular state of that system over a finite period of observation. Further research may demonstrate that the sensitivity of error X reveals useful information about a system, such as identifying features of spatial pattern. However, while this may be useful for other applications, it is unsuitable for estimating CLSs of complex systems.

In contrast, the Pascual and Levin (1999) method provided interpretable spectra and thus clearer estimates of the CLS with increasing model complexity. The prediction r^2 spectra from all model systems were consistently interpretable and the CLS estimates of a

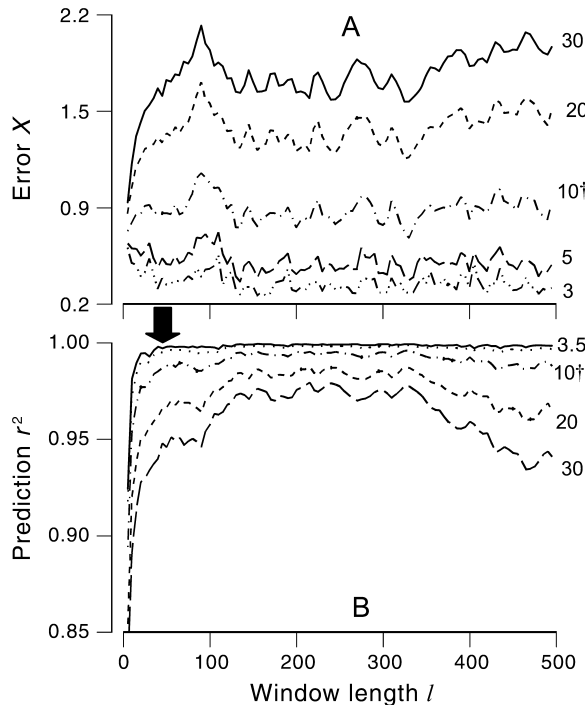


FIG. 8. The effect of increasing numbers of k nearest neighbors (k is given to the right of the curves) for the dominant species group (filamentous turf and coralline algae) of the coral reef model system, based on (A) error X and (B) prediction r^2 . A dagger (\dagger) denotes the k chosen for use in all other model runs. The solid arrow indicates the estimated CLS.

particular system were not sensitive to changes in parameters for attractor reconstruction, or to varying initial conditions. Thus, prediction r^2 spectra using the approach of Pascual and Levin (1999) appears to be a robust means of estimating CLSs of complex ecological systems. We suggest that this metric is likely to be useful and reliable in estimating CLSs of natural ecosystems.

Estimating CLSs: alternative methods without long time series

While Pascual and Levin's (1999) approach to detecting CLSs accommodates oscillatory behavior in dynamic systems and proves reasonably robust over a range of complexities of model ecosystems, the technique requires long time series of data which are not attainable for most natural systems. Our two alternative approaches have considerable potential for overcoming this problem in estimating the CLSs of real ecosystems. These methods detect primary CLSs similar to those determined with the original long time series method (Table 1), but require a maximum of only three to four spatially resolved landscapes instead of thousands.

For the remainder of the discussion, reference to CLSs estimated from the long time series method refers

to those derived using the original method of Pascual and Levin (1999).

CLSs from short time series.—The short time series method uses variability through time as the basis to derive CLS estimates, but employs a unique approach of sampling over the entire landscape at every time step. This generates sufficient data points for nonlinear time series analysis, but removes the requirement for a long series of sequential observations. The underlying motivation of this work is to develop robust techniques for application to real ecosystems, and the requirement of only three or four sampling occasions is realistic for many applications in ecology. Results for the three-species model indicate that the technique generates sensible CLS estimates and is robust to the number of time steps considered. Our more general investigations of the three-, six- and 12-species and coral reef models also indicate that CLS estimates are robust to both the length of the time step and to the number of time steps sampled (detailed results not provided here).

Another advantage of using information from only a small number of time steps is that ecological systems are likely to be essentially stationary in time over the sampling period, whereas the same may not be true of a system's attractor over a longer time period. Furthermore, the short time series approach can be used to detect temporal nonstationarity as a change in the CLS over time (e.g., Fig. 13C, D). Unlike simpler measures such as changes in species' abundances (which occur without any shift in the attractor of a system), a change in the CLS through time indicated from short time series analysis reflects a shift in the underlying dynamics (i.e., the attractor). Note that, in the six-species system for which we demonstrated the shift in the attractor using this approach, the network interaction topology remained constant over time. Thus, the short time series approach could potentially be used to detect fundamental change in natural systems, where a shifting attractor is reflected as a changing CLS. Applied examples of shifting attractors may include systems affected by pollutants or impacted by foreign species.

Similar to the assumption of temporal stationarity that underpins the long time series method, is the assumption of spatial stationarity which applies to our short time series method. This assumption, that species sampled on different parts of the spatial landscape are on the same attractor, cannot be tested by simply examining whether community composition is similar in all areas of space. Different points on the same attractor can, of course, reflect states of vastly dissimilar community structure, for example as arises in simple stable limit cycles. However, in the same way that we detected potential temporal non-stationarity in the attractor by determining the CLS at two different points in time using the short time series method, so could one determine the CLS at two regions in space. If the CLSs are different in different regions of the landscape, then

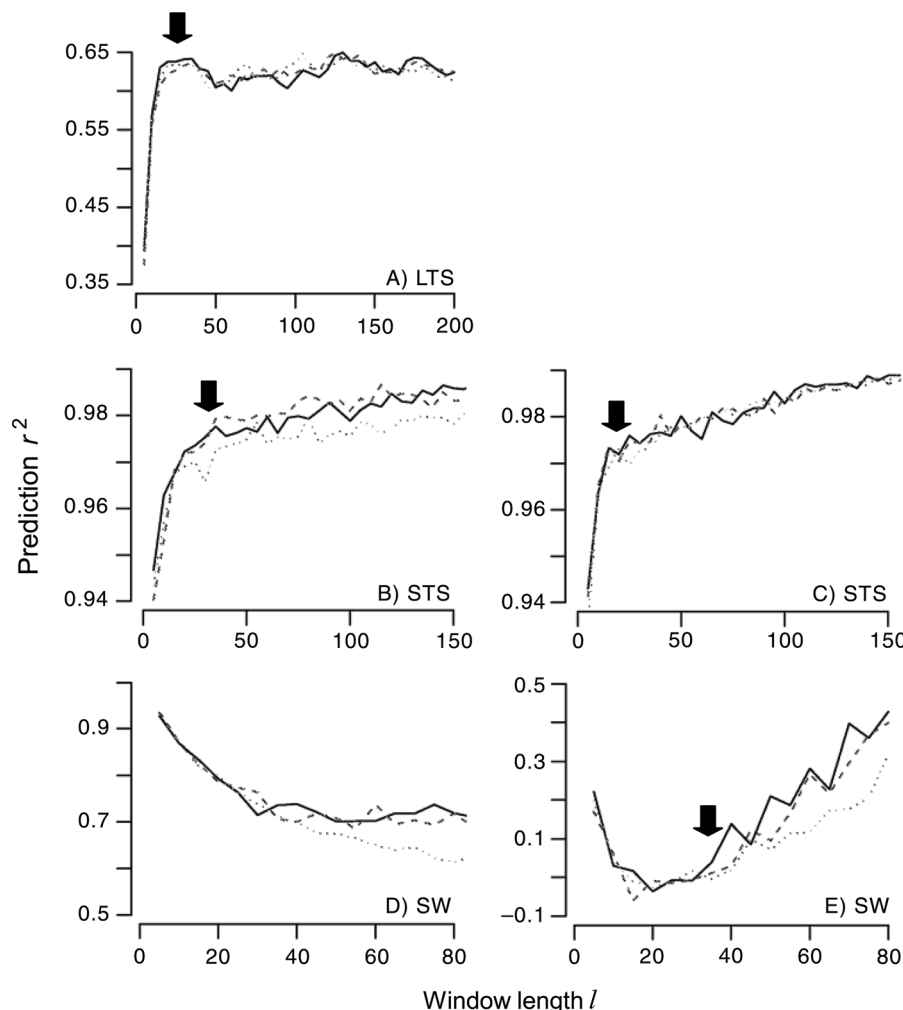


FIG. 9. Prediction r^2 as a function of window length l for the three-species system derived using the three methods. Curves for the three species in the model are shown as different lines. Solid arrows indicate estimated CLSs where curves are interpretable. (A) Long time series analysis of 10 000 time steps with time delay $\tau = 9$ and embedding dimension $d_e = 6$; the first 200 time steps are discarded in the analysis. (B and C) Short time series analysis with $\tau = 1$: (B) spectra derived from analysis of three consecutive time steps (498–500; $d_e = 2$); (C) spectra derived from analysis of four consecutive time steps (497–500; $d_e = 3$). (D and E) Analysis using sliding windows: (D) $\tau = 0.2 \times$ window length l , $d_e = 5$; (E) $\tau = 0.8 \times$ window length l , $d_e = 5$. Parameter choices were based on mutual information and false nearest neighbors techniques. For all analyses, landscape size is 500×500 cells, and the number of nearest neighbors $k = 10$. Key to abbreviations: LTS, long time series analysis; STS, short time series analysis; SW, sliding window analysis.

the attractor cannot be assumed to be spatially stationary across that space. Nevertheless, given the relatively small magnitude of the CLSs and that it is not necessary to sample over large spatial areas to estimate them, we expect that the assumption should hold true provided that a landscape is not sampled across strong environmental gradients or environmental discontinuities likely to realize dissimilar community dynamics.

CLSs could not be identified consistently for the acroporid and faviid corals in the coral reef model based on analysis of short time series. Changing the number of time steps and/or the length of time steps used for attractor reconstruction did not significantly change the shape of the curves. Since both of these coral groups

occur at relatively low densities in this system, the dynamic signal for these groups over short time series may be inadequate for attractor reconstruction and for the determinism test.

CLSs from sliding windows.—Our sliding window method derives CLS estimates from spatial data obtained at a single sampling occasion, using spatial sampling similar to that of others (Allain and Cloitre 1991, Plotnick et al. 1996, Marcos-Nikolaus et al. 2002). The approach replaces temporal variability with spatial variability in reconstructing the system's attractor, but this introduces several complications. In particular, it is clear that “delays” of a fixed number of cells (or pixels) are inappropriate for this method. Consideration

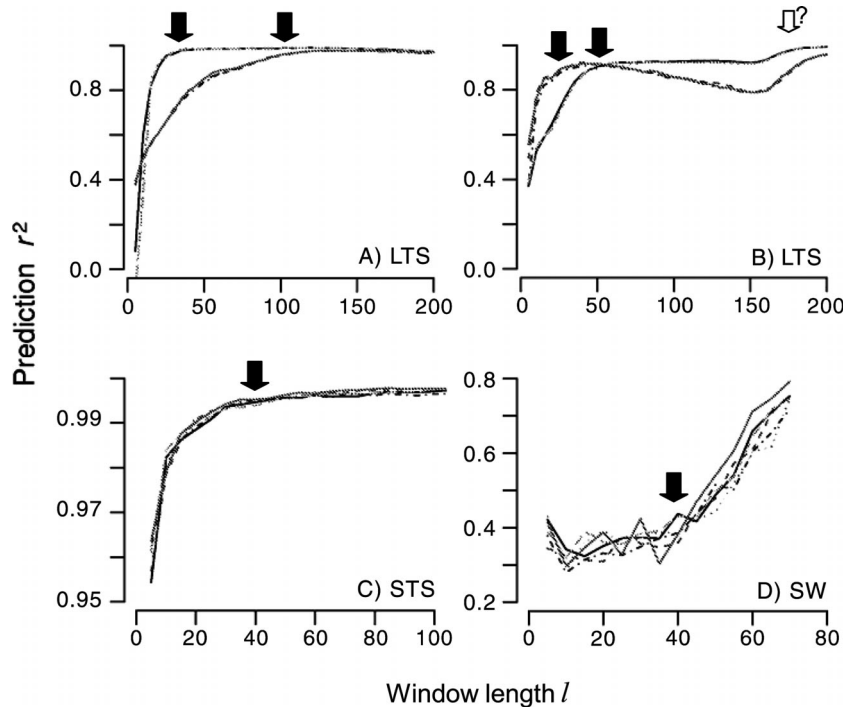


FIG. 10. Prediction r^2 as a function of window length l for the six-species system derived using the three methods. Curves for the six species in the model are shown as different lines. Solid arrows indicate estimated CLSs where curves are interpretable, while open arrows with question marks indicate ambiguous CLSs where interpretation is unclear. (A and B) Long time series analysis of two independent model runs with different initial random configurations of recruits with $\tau = 12$, $d_e = 6$. The analyzed time series contains 10 000 time steps, with the first 200 time steps discarded. (C) Short time series analysis with $\tau = 1$. This analysis is based on three consecutive time steps (498–500; $d_e = 2$). (D) Sliding window analysis of the landscape with $\tau = 0.8 \times \text{window length } l$, and $d_e = 5$. Parameter choices were based on mutual information and false nearest neighbors techniques. For all analyses, landscape size is 500×500 cells and $k = 10$. Key to abbreviations: LTS, long time series analysis; STS, short time series analysis; SW, sliding window analysis.

of prediction r^2 spectra produced from landscapes of independent random pixels (i.e., the null case of no spatial pattern) suggests that delays that are proportional to window size are most appropriate. Proportional delays ensure that the overlap of coordinates is constant for all window sizes (see Appendix D).

For all four model systems, when a delay of $0.8 \times \text{window length } l$ was used (i.e., the delay suggested by the scaled mutual information plot), we obtained CLSs similar to those estimated from the long time series analyses. For sliding window analysis, the CLS was taken as the point of increase of the (inverted) prediction r^2 curve. The inverted shape of prediction r^2 curves using this approach indicates that prediction is good at very small spatial scales, but then declines with increasing window size. The CLS is the point where prediction r^2 then begins to improve again, thereafter increasing with window size. This relationship can be explained in terms of spatial patterns that arise on the model landscapes. At small scales of observation below the size of individual colonies, prediction is good because there is a high likelihood that successive windows in any one “sliding series” will be the same species. At slightly larger scales prediction will be

poorer, because more than one species will occur in sampling windows but the scale of observation will be insufficient to capture the community dynamic among colonies. As window size continues to increase, a point is reached where the spatial pattern becomes more predictable, reflecting the deterministic dynamic among colonies (Johnson 1997, Johnson and Seinen 2002). At this point there is a rise in prediction r^2 , indicating the CLS.

There are some constraints to the general application of the sliding window technique. Although the approach dramatically reduces data requirements compared with the original method based on long time series, the amount of spatial data necessary for the method can be large. Depending on the magnitude of the delay, the maximum possible window size for sliding window analysis may be much smaller than the landscape size. It is therefore necessary to select a landscape size that ensures that the maximum possible window size is larger than the anticipated CLS. Because replication occurs in space, landscapes must generally be larger than those used in the long time series method, where replication occurs by sampling through time. Similarly, landscapes must also be large when using

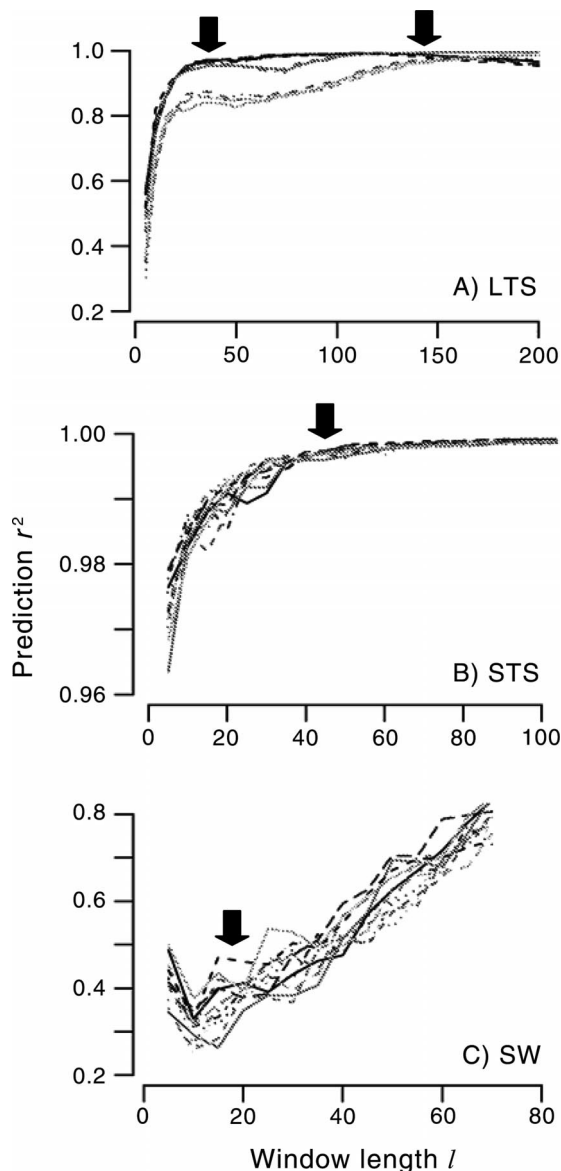


FIG. 11. Prediction r^2 as a function of window length l for the 12-species system derived using the three methods. Curves for the 12 species in the model are shown as different lines. Solid arrows indicate estimated CLSs. (A) Long time series analysis with $\tau = 12$ and $d_e = 6$. The analyzed time series contains 10 000 time steps, with the first 200 time steps discarded. (B) Short time series analysis with $\tau = 1$. This analysis is based on three consecutive time steps (498–500; $d_e = 2$). (C) Sliding window analysis of the landscape with $\tau = 0.8 \times$ window length l , $d_e = 6$. Parameter choices were based on mutual information and false nearest neighbors techniques. For all analyses, landscape size is 500×500 cells and $k = 10$. Key to abbreviations: LTS, long time series analysis; STS, short time series analysis; SW, sliding window analysis.

the short time series method. Thus, when applying these spatial techniques to natural systems, some prior understanding of the general scales of dynamical processes is required to ensure that the maximum sampling scale is large enough to include its CLSs. A second

consideration is that, while the shape of the prediction r^2 curve in the null case of no spatial pattern is acceptable to allow interpretation of CLSs where they do arise, we cannot verify mathematically that this is the expected shape (Appendix C). The reason for sensitivity of CLS estimates to the value selected for the proportional delay is also unclear. However, despite these caveats, the important finding is that using the proportional delay indicated by the minimum mutual information produces a clearly interpretable prediction r^2 spectrum which indicates a CLS similar to that estimated from analysis of a long time series of the same system (Table 1).

Multiple length scales

While different scales will undoubtedly exist that reflect other properties of ecosystems, here we sought the natural scales at which to optimally observe a system's dynamics. It was not surprising that we detected multiple scales in the more spatially complex systems. Several length scales were detected in the complex model systems, either because different species groups indicated different CLSs or a single species indicated multiple CLSs. The mechanisms underpinning these two phenomena are different. Different species in the same system can indicate different scales when the connectivity between groups of species in patches is low. The prediction r^2 spectrum of a single species can also display several length scales, but with the larger scales indicative of that species' emergent dynamics. We discuss each scenario.

Multiple length scales indicated by different species.—In theory, any species in a system can be observed to indicate the length scale of that system, irrespective of its life history or dynamical behavior. However, biologists might intuitively suggest that species with dissimilar life history parameters could be expected to yield dissimilar length scales for the same system. Indeed, our results for the more complex models suggest that estimates of CLS_p show some dependence on the species used for attractor reconstruction. Analyses based on both short and long time series could indicate different scales for different species, but the phenomenon arose more readily in analyses based on long time series.

In long time series analyses, scale differences among species occurred in systems that were strongly spatially self-organizing at several spatial scales, as in the six- and 12-species systems. Notably, these differences were not related to abundances. For example, in the 12-species model the curves for the species form three distinct groups, each with a different estimate of CLS_p (Fig. 5). While all species in this model are topologically equivalent with respect to network structure (each is able to overgrow three others in a symmetrical network), the system spatially self-organizes into three distinct patch types, with each patch type containing four species. Because species within patches are more

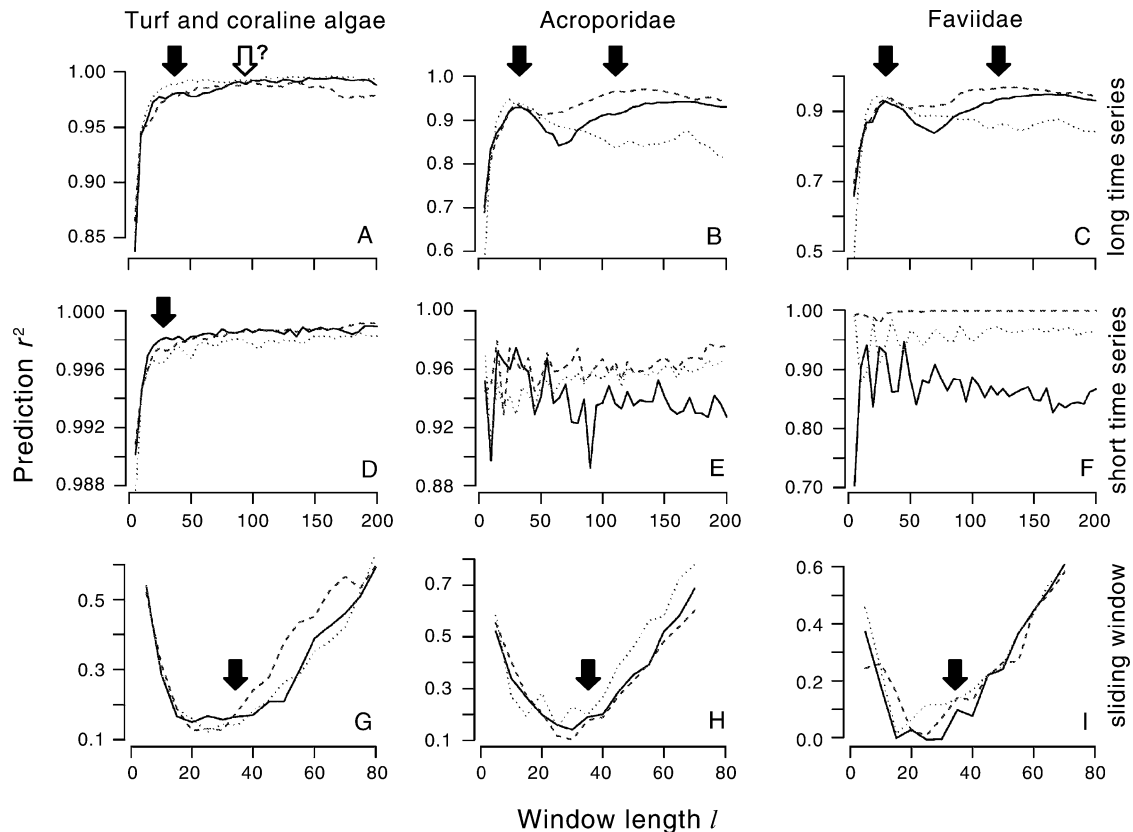


FIG. 12. Prediction r^2 as a function of window length l derived using the three methods applied to cover of three physiognomic groups in the coral reef system: (A, D, G) turf and coralline algae, (B, E, H) corymbose and digitate Acroporidae, and (C, F, I) Faviidae. Curves for three model runs with different initial random configurations of recruits are represented as different lines. Solid arrows indicate estimated CLSs where curves are interpretable, while open arrows with question marks indicate ambiguous CLSs where interpretation is unclear. (A, B, C) Long time series analysis with (A) $\tau = 100$, $d_e = 5$; (B) $\tau = 75$, $d_e = 6$; and (C) $\tau = 100$, $d_e = 6$. The analyzed time series contains 10 000 time steps, with the first 200 time steps discarded. Landscape size is 300×300 cells. (D, E, F) Short time series analyses with $\tau = 1$. These analyses are based on four consecutive time steps (498–500; $d_e = 3$). (G, H, I) Sliding window analysis of the landscape with $\tau = 0.8 \times$ window length l . For (G) $d_e = 5$, and for (H and I) $d_e = 6$. Parameter choices were based on mutual information and false nearest neighbors techniques. For D–I, landscape size is 500×500 cells. For all analyses, $k = 10$.

TABLE 1. Summary of CLS (Pascual and Levin 1999) estimates from long time series (LTS), short time series (STS) and sliding window (SW) analyses for the three-species, six-species, 12-species, and coral reef systems.

System	Analysis	Scale									
		20	40	60	80	100	120	140	160	180	200
Three species	LTS										
	STS										
	SW										
Six species†	LTS										
	STS										
	SW										
12 species	LTS										
	STS										
	SW										
Coral‡	LTS										
	STS										
	SW										

† CLS ranges indicated are for the nonstationary case.

‡ CLS range estimated using short time series analysis is taken for turf and coralline algae only.

likely to interact with each other than with species from other patches given their spatial proximity, the realized connectivity on a spatially organized landscape is higher among species within patches than it is among species between patches. This occurs even though, in this example, the interaction network defines that each species interacts equivalently with others both inside and outside a patch type.

Not surprisingly, highly connected species with tightly coupled dynamics (i.e., species within patches) indicate similar length scales, while species whose dynamics are more weakly linked (because they are spatially separated in different patches) can manifest dissimilar length scales, even though they have identical “life history” attributes. In our systems where different length scales arise for different species, differential connectivity among species arises through spatial self-organizing. We anticipate that other factors which influence the connectivity among species, such as the topology of food webs in which some groups of species are tightly coupled trophically while others are poorly connected (e.g., O’Neill et al. 1986, Johnson et al. 1995), will have a similar effect. Because the approach is based on reconstruction of deterministic dynamics, it should not be surprising that species that are weakly linked dynamically can indicate different length scales for their different behaviors. Conversely, species that are highly connected in a system, as in our three-species model, all indicate very similar estimates of CLS_p .

In the coral reef model, the different physiognomic groups provide similar CLS_p estimates despite notable differences among groups in their life history parameters such as rates of recruitment, growth, and mortality. Marked differences in CLS_p estimates among “species” (in reality, guilds) in this model do not arise because they do not self-organize into distinct patches. In one sense, this system is maximally connected in that each species competes with all others for space and there are no standoffs (Johnson and Seinen 2002). However, the species are not identically connected because they have different interaction strengths and neighbor-specific growth rates.

In short time series analyses using both the six- and 12-species models, the spectra of species from different patches also separated, revealing different length scales among species. With this method, the different CLS s likely reflect the connectivity of spatial patches captured in the set of three to four landscapes that are used to reconstruct the attractor. However, in general, spectra produced from long time series tended to more clearly differentiate among weakly connected species.

In summary, because CLS estimates reflect the strength of dynamical connectivity among species, we predict that for complex real systems different species or functional groups that are loosely connected, either as a result of spatial separation or weak direct interactions, may indicate dissimilar length scales. This will

arise whether the long time series method or the short time series method is used.

Multiple length scales indicated by a single species.—For complex model systems, analyses based on long time series from a single species may detect several different length scales, reflecting the different dynamics within the system. For example, the prediction r^2 spectrum of a single group in the six-species system sometimes displays several critical points (Fig. 4B), which we interpret as multiple length scales. The smallest length scale (the first peak in the curve) is consistent among runs, and is also consistently indicated by our alternative short time series and sliding window methods. This “primary scale” is the scale at which the local dynamic is best predicted, and it reflects the scale of interaction between colonies of different species within patches. Thus, for any one species, we expect at least one CLS , which is influenced by colony size, to reflect the nature of local interactions between species within patches. In the models we examined in which both distinct colonies and clearly differentiated groups of colonies (i.e., patches) formed, the primary CLS is larger than the mean colony size but smaller than the patch size. In real ecological systems, we might expect the primary CLS to indicate the most appropriate scale for monitoring system dynamics, and therefore the scale that most efficiently identifies meaningful trends in species abundances.

Scales larger than the primary scale likely reflect the system’s emergent dynamics, which may include the emergence of and interactions among patches, and non-stationary attractors. Because emergent dynamics are highly variable among runs (e.g., the size and shape of patches vary), length scales larger than the primary scale tend to be more variable among runs (e.g., Fig. 5B, D). Thus, while any single run may demonstrate several length scales, the average of several prediction r^2 curves for a given species is likely to indicate only a single CLS (e.g., Fig. 3B). The widening of the 95% confidence intervals around the average curve reflects that the variable secondary peaks are absorbed as noise (Fig. 3B).

While the short time series method occasionally identified secondary CLS s, it is clear that for analyses based on a single species, the short time series and sliding window methods are less likely to identify additional CLS s larger than the primary CLS than is the method based on long time series. This result may indicate that emergent dynamics are more difficult to detect from spatial pattern than from time series. When using the short time series or sliding window method, if it is important to detect secondary length scales reflecting emergent dynamics, it may be necessary to sample many windows that collectively cover areas much larger than the primary length scale to more comprehensively sample the attractor. The tradeoff in this approach is that, as the distance over which a landscape is sampled increases, more care is likely to be required

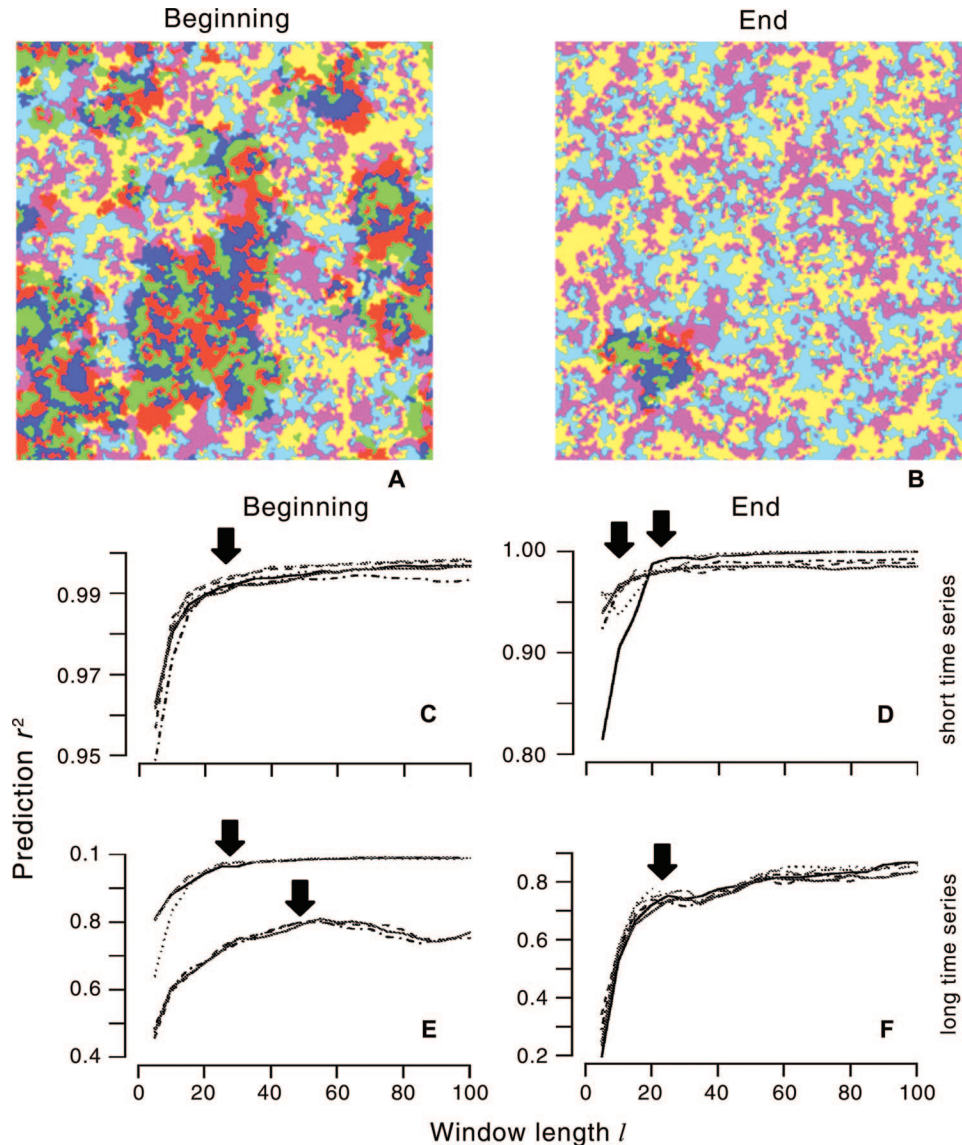


FIG. 13. Comparison of prediction r^2 spectra derived from short and long time series approaches for the six-species system. Arrows identify estimated CLSs. In (C) and (D), short time series analysis uses time steps from the start (A) and end (B) of a 5000-time step time series. These analyses are based on four consecutive time steps ($d_e = 3$) with $\tau = 1$: (C) time steps 201–204; (D) time steps 4997–5000. Divergence in species abundances by the end of the time series is reflected in the two distinct groupings of prediction r^2 curves in (D). (E and F) Long time series analysis for a nonstationary and a stationary system, respectively (see *Results: Short time series analysis* for explanation). The analyzed time series in (E) and (F) contains 5000 data points, sampled at each time step. For (E), $\tau = 12$, $d_e = 6$; for (F), $\tau = 19$, $d_e = 6$. Parameter choices were based on mutual information and false nearest neighbors techniques. For all analyses, landscape size is 200×200 cells, and $k = 10$.

to ensure that all samples are represented by the same attractor.

Do CLSs have a future in ecology?

Both the short time series and sliding window techniques we describe offer workable alternatives to the existing approach based on analyses of long time series (Pascual and Levin 1999) for estimating the primary CLS of dynamical systems. The distinct advantage of

these new approaches is that their data requirements can be met for natural ecosystems. This is an important step towards the goal of estimating CLSs of real ecosystems. Of our two alternatives, the short time series method seems to be the most robust and interpretable, combining both temporal and spatial data to reconstruct the attractor.

Several questions arise with regards to the application of the new CLS methods to natural ecosystems.

Do these new methods unambiguously identify the primary CLS of the system? We have shown that the short time series method and sliding windows method can be used to readily detect primary CLSs virtually identical to those identified by the original long time series method, in some cases with less ambiguity (Table 1).

Another important issue is whether a system has only a single length scale at which it is optimal to observe dynamics, as suggested by the current theory, or whether there can be many. Most ecologists would argue that any ecological system is likely to manifest several length scales (Levin 2000, Bishop et al. 2002), and it is therefore important for the techniques to have the potential to identify them. Using the long time series method, different length scales are indicated by different species within the same system if their dynamics are weakly linked (e.g., Figs. 4B, 5B, D). Using our new method based on a short series of three to four time steps, species in different patches with weakly coupled dynamics also indicate different CLSs in the six- (Fig. 13D) and 12-species (not shown) model systems. Thus, when applied to a natural system in which subcomponents are weakly linked, this technique should have the ability to indicate that the partially decoupled dynamics among subsets of weakly linked species are best observed at different spatial scales. Similarly, these approaches to estimating length scales can identify secondary scales that indicate emergent dynamics. Finally, we note that detection of natural scales for optimal observation of ecosystem dynamics does not preclude that there may be other natural scales that reveal other ecologically meaningful properties of real ecosystems.

Our overall conclusion is that the development of new techniques that can be realistically applied in ecology to produce prediction r^2 spectra for complex oscillating dynamical systems is useful progress towards objectively defining appropriate scales for observing natural ecological systems. The next challenge is to use the modified techniques to assess what CLSs inform us about natural systems, and to evaluate their utility in providing objective estimates for scaling issues in applied ecology.

ACKNOWLEDGMENTS

We are grateful to Mercedes Pascual for providing us with a data set for verification of the prediction algorithms. We also thank Simon Levin, Matthew Keeling, Mark Denny, and an anonymous reviewer for valuable comments on the manuscript, and Piers Dunstan for discussion of several critical aspects of the research. This research was supported by an ARC Discovery Grant, and funds from the Tasmanian Aquaculture and Fisheries Institute and the School of Zoology, University of Tasmania, all awarded to C. Johnson. R. Habeeb was supported by a Thomas Crawford Memorial Scholarship.

LITERATURE CITED

- Abarbanel, H. 1996. Analysis of observed chaotic data. Springer-Verlag, New York, New York, USA.
- Allain, C., and M. Cloitre. 1991. Characterizing the lacunarity of random and deterministic fractal sets. *Physical Review A* **44**:3552–3558.
- Bishop, M. J., A. J. Underwood, and P. Archambault. 2002. Sewage and environmental impacts on rocky shores: necessity of identifying relevant spatial scales. *Marine Ecology Progress Series* **236**:121–128.
- Buzug, T., and G. Pfister. 1992. Comparison of algorithms calculating optimal embedding parameters for delay time coordinates. *Physica D* **58**:127–137.
- Carlile, D., J. Skalski, J. Batker, J. Thomas, and V. Cullinan. 1989. Determination of ecological scale. *Landscape Ecology* **2**:203–213.
- Casdagli, M. 1989. Nonlinear prediction of chaotic time series. *Physica D* **35**:335–356.
- De Roos, A., E. McCauley, and W. Wilson. 1991. Mobility versus density-limited predator-prey dynamics on different spatial scales. *Proceedings of the Royal Society of London B* **246**:117–122.
- Dieckmann, U., R. Law, and J. A. J. Metz, editors. 2000. The geometry of ecological interactions: simplifying spatial complexity. Cambridge University Press, Cambridge, UK.
- Dunstan, P. K., and C. R. Johnson. 2005. Predicting global dynamics from local interactions: individual-based models predict complex features of marine epibenthic communities. *Ecological Modelling*, **186**:221–233.
- Durrett, R., and S. A. Levin. 2000. Lessons on pattern formation from planet WATOR. *Journal of Theoretical Biology* **205**:201–214.
- Ellner, S., and P. Turchin. 1995. Chaos in a noisy world: new methods and evidence from time-series analysis. *American Naturalist* **145**:343–375.
- Farmer, J. 1982. Chaotic attractors of an infinite-dimensional dynamical system. *Physica D* **4**:366–393.
- Greig-Smith, P. 1952. The use of random and contiguous quadrats in the study of the structure of plant communities. *Annals of Botany* **16**:293–316.
- Hastings, A., C. L. Hom, S. Ellner, P. Turchin, and H. C. J. Godfray. 1993. Chaos in ecology: is mother nature a strange attractor? *Annual Review of Ecology and Systematics* **24**:1–33.
- Johnson, C. R. 1997. Self-organising in spatial competition systems. Pages 245–263 *in* N. Klomp and I. Lunt, editors. *Frontiers in ecology: building the links*. Elsevier, Oxford, UK.
- Johnson, C. R., D. Klumpp, J. Field, and R. Bradbury. 1995. Carbon flux on coral reefs: effects of large shifts in community structure. *Marine Ecology Progress Series* **126**:123–143.
- Johnson, C. R., and K. H. Mann. 1988. Diversity, patterns of adaptation, and stability of Nova Scotian kelp beds. *Ecological Monographs* **58**:129–154.
- Johnson, C. R., and I. Seinen. 2002. Selection for restraint in competitive ability in spatial competition systems. *Proceedings of the Royal Society of London B* **269**:655–663.
- Kantz, H., and T. Schreiber. 1997. Nonlinear time series analysis. Cambridge University Press, Cambridge, UK.
- Kaplan, D., and L. Glass. 1995. Understanding nonlinear dynamics. Springer-Verlag, New York, New York, USA.
- Keeling, M. J., I. Mezic, R. Hendry, J. Mcglade, and D. Rand. 1997. Characteristic length scales of spatial models in ecology via fluctuation analysis. *Philosophical Transactions of the Royal Society of London B* **352**:1589–1601.
- Kennel, M. B., R. Brown, and H. D. I. Abarbanel. 1994. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physics Review A* **45**:3403–3411.
- Kershaw, K. A. 1957. The use of cover and frequency in the detection of pattern in plant communities. *Ecology* **38**:291–299.

- Levin, S. A. 1992. The problem of pattern and scale in ecology. *Ecology* **73**:1943–1967.
- Levin, S. A. 2000. Multiple scales and the maintenance of biodiversity. *Ecosystems* **3**:498–506.
- Levin, S. A., B. Grenfell, A. Hastings, and A. S. Perelson. 1997. Mathematical and computational challenges in population biology and ecosystems science. *Science* **275**:334–343.
- Liebert, W., and H. Schuster. 1989. Proper choice of the time delay for the analysis of chaotic time series. *Physics Letters A* **142**:107–111.
- Little, S., S. Ellner, M. Pascual, M. Neubert, D. Kaplan, T. Sauer, H. Caswell, and A. Solow. 1996. Detecting nonlinear dynamics in spatio-temporal systems, examples from ecological models. *Physica D* **96**:321–333.
- Marcos-Nikolaus, P., J. M. Martin-Gonzalez, and R. V. Sole. 2002. Spatial forecasting: detecting determinism from single snapshots. *International Journal of Bifurcation and Chaos* **12**:369–376.
- Molofsky, J., J. Bever, J. Antonovics, and T. Newmaan. 2002. Negative frequency dependence and the importance of spatial scale. *Ecology* **83**:21–27.
- Nichols, J. M., and J. D. Nichols. 2001. Attractor reconstruction for non-linear systems: a methodological note. *Mathematical Biosciences* **171**:21–32.
- O'Neill, R. V., D. L. DeAngelis, J. B. Waide, and T. F. H. Allen. 1986. A hierarchical concept of ecosystems. Princeton University Press, Princeton, New Jersey, USA.
- Pascual, M., and S. Ellner. 2000. Linking ecological patterns to environmental forcing via nonlinear time series models. *Ecology* **81**:2767–2780.
- Pascual, M., and S. A. Levin. 1999. From individuals to population densities: searching for the intermediate scale of nontrivial determinism. *Ecology* **80**:2225–2236.
- Plotnick, R. E., R. H. Gardner, W. W. Hargrove, K. Prestegard, and M. Perlmutter. 1996. Lacunarity analysis: a general technique for the analysis of spatial patterns. *Physical Review E* **53**:5461–5468.
- Rand, D. 1994. Measuring and characterizing spatial patterns, dynamics and chaos in spatially extended dynamical systems and ecologies. *Philosophical Transactions of the Royal Society of London A* **348**:497–514.
- Rand, D., and H. Wilson. 1995. Using spatio-temporal chaos and intermediate-scale determinism to quantify spatially extended ecosystems. *Proceedings of the Royal Society of London B* **259**:111–117.
- Schneider, D. C. 1994. Quantitative ecology: spatial and temporal scaling. Academic Press, San Diego.
- Schreiber, T. 1995. Efficient neighbor searching in nonlinear time series analysis. *International Journal of Bifurcation and Chaos* **5**:349–358.
- Schreiber, T. 1999. Interdisciplinary application of nonlinear time series methods. *Physics Reports* **308**:2–64.
- Sole, R. V., and J. Bascompte. 1995. Measuring chaos from spatial information. *Journal of Theoretical Biology* **175**:139–147.
- Sugihara, G., B. Grenfell, and R. M. May. 1990. Distinguishing error from chaos in ecological time series. *Philosophical Transactions of the Royal Society of London B* **330**:235–251.
- Takens, F. 1981. Detecting strange attractors in turbulence. Pages 366–381 in D. Rand and L. Young, editors. *Dynamical systems and turbulence*, Warwick 1980. Lecture notes in mathematics. Springer-Verlag, New York, New York, USA.
- Tilman, D., and P. Kareiva, editors. 1997. *Spatial ecology: the role of space in population dynamics and interspecific interactions*. Princeton University Press, Princeton, New Jersey, USA.
- Turner, S., R. V. O'Neill, W. Conley, M. Conley, and H. Humphries. 1991. Pattern and scale: Statistics for landscape ecology. Pages 17–47 in S. J. Turner and R. H. Gardner, editors. *Quantitative methods in landscape ecology*. Springer Verlag, New York, New York, USA.
- Tyre, A. J., H. P. Possingham, and C. M. Bull. 1997. Characteristic scales in ecology: fact, fiction or futility. Pages 233–243 in N. Klomp and I. Lunt, editors. *Frontiers in ecology*. Elsevier Science, New York, New York, USA.
- Wiens, J. 1989. Spatial scaling in ecology. *Functional Ecology* **3**:385–397.
- Wilson, H. B., and M. J. Keeling. 2000. Spatial scales and low dimensional deterministic dynamics. Pages 209–226 in U. Dieckmann, R. Law, and J. A. J. Metz, editors. *The geometry of ecological interactions: simplifying spatial complexity*. Cambridge University Press, Cambridge, UK.

APPENDIX A

A figure showing species' trajectories using different sized windows of observation is available in ESA's Electronic Data Archive: *Ecological Archives* M075-018-A1.

APPENDIX B

A discussion of attractor reconstruction using nonlinear time series analysis is available in ESA's Electronic Data Archive: *Ecological Archives* M075-018-A2.

APPENDIX C

A discussion of prediction r^2 and error X in the null case of no spatial pattern is available in ESA's Electronic Data Archive: *Ecological Archives* M075-018-A3.

APPENDIX D

A discussion of the evaluation of fixed and proportional delays is available in ESA's Electronic Data Archive: *Ecological Archives* M075-018-A4.

SUPPLEMENT

The COMPETE software (a menu-driven individual-based spatial modeling program) and user manual are available in ESA's Electronic Data Archive: *Ecological Archives* M075-018-S1.